NCGAS: National Resources for Computationally Intensive Bioinformatics

Carrie Ganote, Barb Hallock, Sheri Sanders, Tom Doak

PERVASIVE TECHNOLOGY INSTITUTE
INDIANA UNIVERSITY

BRIDGES
A PITTSBURGH SUPERCOMPUTING CENTER RESOURCE

NSF

Genomics as a data-intensive science

Bioinformatics—especially genomics and new imaging methods—has taken its place as a Big Data science. For example, next generation sequencing technologies have increased the throughput of sequencing while decreasing the cost dramatically. The plummeting cost has led to a proliferation of techniques to explore phenomena such as transcriptomics, methylation and metagenomics.

Why NCGAS does Genomics?

- Genomics, transcriptomics, etc. 'omics have and are revolutionizing everything from ecology to cancer treatment.
- Genomics is computationally demanding
- There are many new researchers, whose specialties lie elsewhere

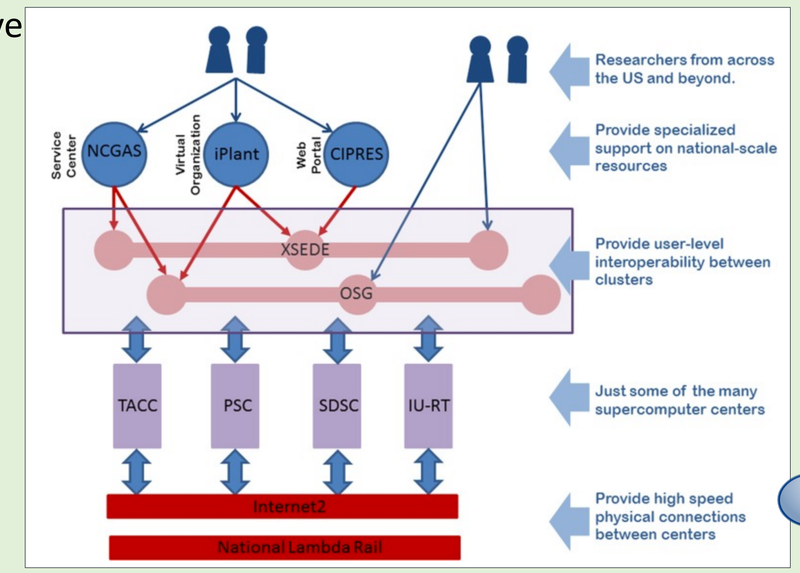
IU has a rich history in genomics, and RT in helping them.

- NSF agreed to fund our services

NCGAS's primary goals:

- Provide bioinformatics expertise to NSF researchers
- Maintain a curated set applications
- Provide access to HPC resources, esp. large memory clusters = Mason, Bridges
- Build Galaxy instances for our software
- Pursue outreach to biologists

NCGAS serves as a entry point to the research information superhighway:



As an ITCR partner:

IU students and faculty have access to Galaxy @ IU using their Indiana University credentials. Even with access to the resources available at IU, the interface between users and computer clusters can be daunting. We provide support to IU affiliates through Galaxy to accomplish their bioinformatics analyses without the need for a degree in computer science.

Click here to enter Galaxy @ IU

Trinity CTAT Galaxy, hosted by Indiana University and the Broad Institute, is a free-to-use public interface for Trinity users

Click here to enter Trinity_CTAT Galaxy

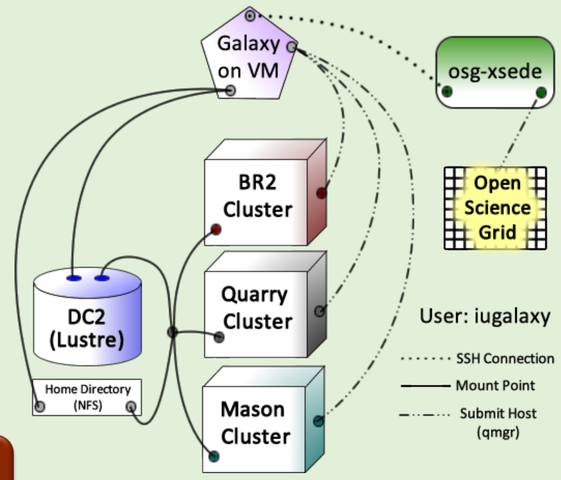
GenePattern is a freely available computational biology open-source software package developed at the Broad Institute of MIT and Harvard, for the analysis of genomic data. NCGAS now hosts a free GenePattern server, with increased computational resources.

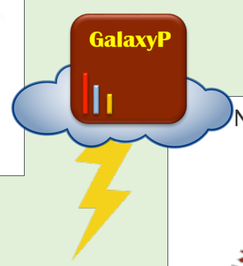
Click here to enter IU's Public GenePattern server

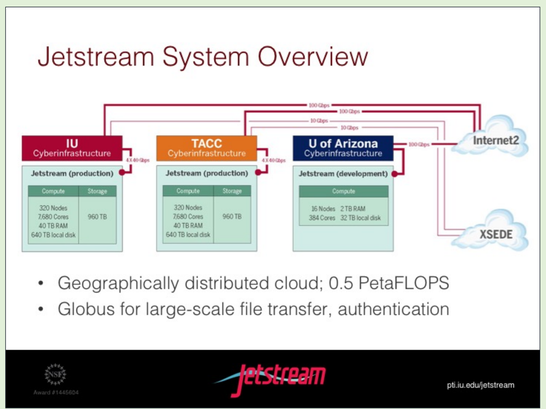
Galaxy Main, hosted by Penn State University and Emory University, is a free-to-use public service that includes hundreds of tools and a server with 250 GB of storage space per user. Our own Galaxy instances are based off of this technology.

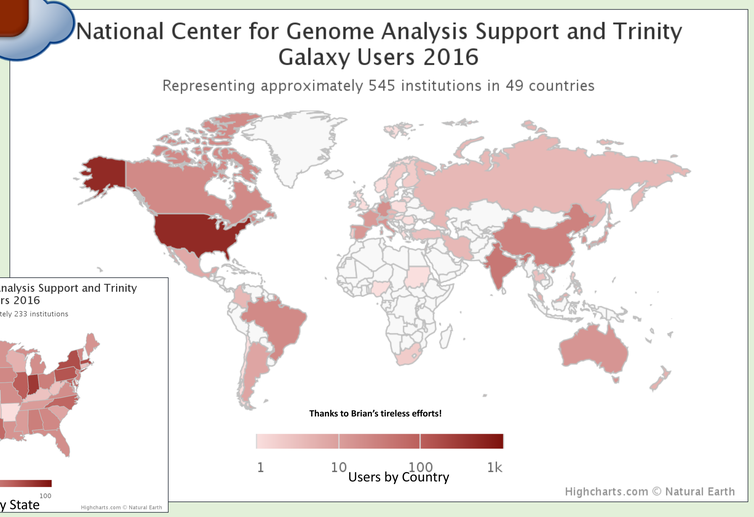
Click here to enter Galaxy Main

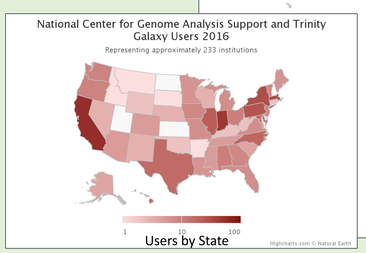
The GALAXY.NCGAS.ORG model











NCGAS.ORG

NSF Awards DBI-1458641 and ABI-1062432